

01059
M. J. Flynn J. N. Gray A. K. Jones K. Lagally
H. Opderbeck G. J. Popek B. Randell
J. H. Saltzer H. R. Wiehle

Operating Systems

An Advanced Course

Edited by
R. Bayer, R. M. Graham, and G. Seegmüller



Springer-Verlag Berlin Heidelberg New York

211 7. 21059
M. J. Flynn J. N. Gray A. K. Jones K. Lagally
H. Opderbeck G. J. Popek B. Randell
J. H. Saltzer H. R. Wiehle

Operating Systems

An Advanced Course

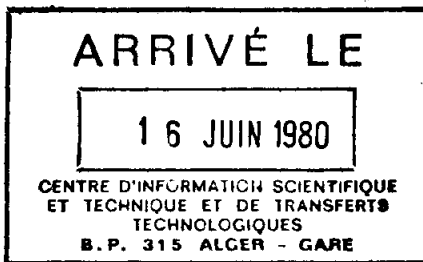
Edited by
R. Bayer, R. M. Graham, and G. Seegmüller



Springer-Verlag
Berlin Heidelberg New York 1979

P R E F A C E

The Advanced Course on Operating Systems was held at the Technical University in Munich from July 28 to August 5, 1977, and was repeated from March 29 to April 6, 1978. The course was organized by the Institute for Informatics of the Technical University Munich and the Leibniz Computing Center of the Bavarian Academy of Sciences, in co-operation with the European Communities, sponsored by the Ministry for Research and Technology of the Federal Republic of Germany.



C o n t e n t s

CHAPTER 1.: INTRODUCTION

R. Bayer R. M. Graham J. H. Saltzer G. Seegmüller	INTRODUCTION	1
--	--------------	---

CHAPTER 2.: MODELS

A. K. Jones	THE OBJECT MODEL: A CONCEPTUAL TOOL FOR STRUCTURING SOFTWARE	7
	1. <i>The Object Model</i>	8
	2. <i>The Object Model Applied to Operating Systems</i>	11
	3. <i>Mechanics of Supporting Type Modules</i>	15
	4. <i>Observation</i>	16
	5. <i>References</i>	16
M. J. Flynn	COMPUTER ORGANIZATION AND ARCHITECTURE	17
	1. <i>Machine Mapping and Well Mapped Machines</i>	19
	2. <i>Name Space - Memory Space</i>	37
	3. <i>Traditional Machine Language Problems and Some Fundamental Concepts</i>	52
	4. <i>Towards Ideal Program Representations</i>	56
	5. <i>Parallel Processor Forms of Computing Systems</i>	81
	<i>References</i>	97

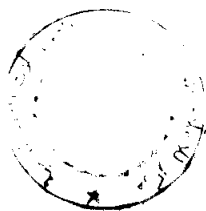
CHAPTER 3.: ISSUES AND RESULTS IN THE DESIGN OF OPERATING SYSTEMS

J. H. Saltzer	NAMING AND BINDING OF OBJECTS	99
	<u>A. Introduction</u>	102
	1. <i>Names in Computer Systems</i>	102
	2. <i>A Model for the Use of Names</i>	104
	3. <i>Problems in the Use of Names</i>	110
	4. <i>Some Examples of Existing Naming Systems</i>	114
	5. <i>The Need for Names with Different Properties</i>	120
	6. <i>Plan of Study</i>	123
	<u>B. An Architecture for Addressing Shared Objects</u>	124
	1. <i>User-Dependent Bindings and Multiple Naming Contexts</i>	129
	2. <i>Larger Contexts and Context Switching</i>	136

5.2.	<i>Authority Lists</i>	241
5.3.	<i>Capability Based Implementation</i>	242
5.3.1.	<i>Extended Object Types</i>	244
5.3.2.	<i>Status</i>	248
6.	<i>Enforcing Information Control Policies</i>	248
7.	<i>References</i>	250
K. Lagally	SYNCHRONIZATION IN A LAYERED SYSTEM	252
1.	<i>Introduction</i>	253
2.	<i>General Concepts</i>	253
2.1.	<i>Synchronization</i>	253
2.2.	<i>Processes and Messages</i>	255
2.3.	<i>Process Hierarchy</i>	257
3.	<i>Implementation Tools</i>	258
3.1.	<i>Semaphores</i>	258
3.2.	<i>Conditional Critical Regions</i>	259
3.3.	<i>Monitors</i>	260
3.4.	<i>Path Expressions</i>	261
3.5.	<i>Object Managers</i>	262
4.	<i>Examples</i>	263
4.1.	<i>Readers and Writers</i>	263
4.1.1.	<i>Semaphores</i>	264
4.1.2.	<i>Conditional Critical Regions</i>	268
4.1.3.	<i>Monitors</i>	269
4.1.4.	<i>Path Expressions</i>	270
4.1.5.	<i>Object Managers</i>	271
4.2.	<i>The Five Dining Philosophers</i>	275
5.	<i>Conclusion</i>	277
6.	<i>References</i>	278
B. Randell	RELIABLE COMPUTING SYSTEMS	282
1.	<i>Introduction</i>	283
2.	<i>Basic Concepts</i>	286
2.1.	<i>Systems and Their Failures</i>	286
2.2.	<i>Errors and Faults</i>	287
3.	<i>Reliability Issues</i>	290
3.1.	<i>Requirements</i>	290
3.2.	<i>Types of Fault</i>	291
3.3.	<i>Fault Intolerance and Fault Tolerance</i>	293
3.4.	<i>Design Fault Tolerance</i>	294
4.	<i>System Structure</i>	296
4.1.	<i>Static Structure</i>	296
4.2.	<i>Dynamic Structure</i>	298
4.3.	<i>Atomic Actions</i>	299
4.4.	<i>Forms of Atomic Action</i>	302
4.5.	<i>Levels of Abstraction</i>	303
4.6.	<i>Faults and Structuring</i>	306
5.	<i>Fault Tolerance Techniques</i>	308
5.1.	<i>Protective Redundancy</i>	308
5.1.1.	<i>Triple Modular Redundancy</i>	309
5.2.	<i>Error Detection</i>	311
5.2.1.	<i>Types of Check</i>	312
5.2.2.	<i>Interface Checking</i>	313
5.3.	<i>Fault Treatment</i>	314
5.4.	<i>Damage Assessment</i>	317
5.5.	<i>Error Recovery</i>	318
5.5.1.	<i>Backward Error Recovery</i>	318

3.5.	<i>Views</i>	409
3.5.1.	<i>Views and Update</i>	411
3.6.	<i>Structure of Data Manager</i>	411
3.7.	<i>A Sample Data Base Design</i>	412
3.8.	<i>Comparison to File Access Method</i>	414
3.9.	<i>Bibliography</i>	414
4.	<i>Data Communications</i>	415
4.1.	<i>Messages, Sessions, and Relationship to Network Manager</i>	415
4.2.	<i>Session Management</i>	417
4.3.	<i>Queues</i>	417
4.4.	<i>Message Recovery</i>	418
4.5.	<i>Response Mode Processing</i>	418
4.5.	<i>Conversations</i>	419
4.6.	<i>Message Mapping</i>	419
4.7.	<i>Topics not Covered</i>	420
4.8.	<i>Bibliography</i>	420
5.	<i>Transaction Management</i>	421
5.1.	<i>Transaction Scheduling</i>	424
5.2.	<i>Distributed Transaction Management</i>	425
5.3.	<i>The Data Management System as a Subsystem</i>	427
5.4.	<i>Exception Handling</i>	428
5.5.	<i>Other Components Within Transaction Management</i>	429
5.6.	<i>Bibliography</i>	429
5.7.	<i>Lock Management</i>	430
5.7.1.	<i>Pros and Cons of Concurrency</i>	430
5.7.2.	<i>Concurrency Problems</i>	431
5.7.3.	<i>Model of Consistency and Lock Protocols</i>	431
5.7.4.	<i>Locking, Transaction Backup and System Recovery</i>	437
5.7.5.	<i>Lower Degrees of Consistency</i>	438
5.7.6.	<i>Lock Granularity</i>	438
5.7.7.	<i>Lock Management Pragmatics</i>	446
5.7.8.	<i>Bibliography</i>	458
5.8.	<i>Recovery Management</i>	459
5.8.1.	<i>Model of Errors</i>	459
5.8.2.	<i>Overview of Recovery Management</i>	460
5.8.3.	<i>Recovery Protocols</i>	462
5.8.4.	<i>Structure of Recovery Manager</i>	472
5.8.5.	<i>Log Management</i>	478
5.8.6.	<i>Examples of a Recovery Routine</i>	480
5.8.7.	<i>Historical Note on Recovery Management</i>	480
5.8.8.	<i>Bibliography</i>	481

H. Opderbeck	COMMON CARRIER PROVIDED NETWORK INTERFACES	482
1.	<i>Introduction</i>	483
2.	<i>Protocoll Characteristics</i>	485
2.1.	<i>Connection Establishment and Clearing</i>	485
2.2.	<i>Error Control</i>	485
2.3.	<i>Flow Control</i>	486
2.4.	<i>Multiplexing</i>	487
2.5.	<i>Synchronisation</i>	488
2.6.	<i>Transparency</i>	489
3.	<i>Terminal Emulation Interface</i>	490
4.	<i>Character Concentration Interface</i>	493
5.	<i>X.25 Interface</i>	495
5.1.	<i>Introduction</i>	495
5.2.	<i>Link Access Procedure</i>	496
5.3.	<i>Packet Level Interface</i>	498



CHAPTER 1: INTRODUCTION

R. Bayer

Technical University Munich
Munich, Germany

R. M. Graham

University of Massachusetts
Amherst, Mass., USA

J. H. Saltzer

Massachusetts Institute of Technology
Cambridge, Mass., USA

G. Seegmüller

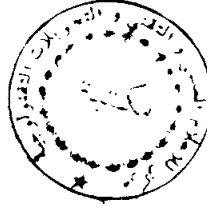
Leibniz Computing Center
of the Bavarian Academy of Sciences
Munich, Germany

Programming systems consisting of

editors, compilers, debuggers, ...
 the operating system,
 the hardware.

Data base systems consisting of

data base managers,
 the operating system,
 the hardware.

Application systems consisting of

application programs,
 the operating system,
 the hardware.

There is also agreement on those aspects that are at the heart of operating systems. In fact, the terms nucleus or kernel are often used for the most essential functions of an operating system. Much of the research and development in operating systems has focused on resource management and the user's interface to this management. Our view of operating systems and the focus of this course is resource management in a very wide sense and the attendant user interface. We shall concentrate on the semantics of this interface, on internal system structure and, to some extent, on hardware architecture.

It is interesting and instructive to look briefly at the history of modern computer systems. In the beginning, computers were small, simple, and free standing. Each individual could use the machine on a one-to-one basis. Generally, there has been an evolution from this state to the current large, complex, multiprogramming, multiprocessor, central systems with virtual memory and many ancillary devices and subsystems. The major trends have been: from one user to many users of the same system; from isolated users to cooperating users; from sequential batch to multiprogramming, to time sharing; and, in both hardware and software, an increase in the degree of concurrency. Most importantly, we see a trend toward increased concern with the management of non-physical resources.

The first computer users always had the entire computer all to themselves for some interval of time. A user always had all the resources. Any resource management facilities provided by an operating (or programming) system were entirely for the user's convenience. As the user community grew it was necessary to insure efficient,

the problem to be solved. The notion of an abstract machine which is available to each user encompasses the essence of this direction of abstraction.

What is the current state of affairs? In a recent workshop the lecturers of this course concluded that the classic problems of physical resource management and concurrency management are well understood, at least to the extent that their implementation is routine and minor enough that operating systems that are satisfactory to the market place are being built. We have chosen to omit from this course any consideration of these problems. Acceptable solutions are widely known. In fact, all of the recent textbooks on operating systems contain extensive discussions of these problems and their solutions. Rather we tried to focus on problems that were less well understood in the past - that are on or near the frontier of the field and that showed significant progress within the last few years. For example, none of the textbooks has an adequate discussion of protection, yet this is one of the most important problems in the design of new operating systems.

Abstractions are based on models. We recognize that models are not only needed to cope with complexity, but ultimately they are needed to verify or validate the correctness and other desired properties of a specific system design. Models for the underlying hardware are the foundation upon which more abstract, general models are built, since they give us insight into the fundamental mechanisms for the final interpretation of a program that is required to produce actual results. In addition, through them we can glimpse a future kind of architecture with many parallel activities, highly distributed.

The object model is the basis for the abstract resource, an object. This very general model is applicable to both software and hardware. It has benefitted from more recent developments in the study of programming languages. This benefit is not incidental. There, the need for careful specification of interfaces with total protection of their implementation has led to the introduction of abstract data types. Objects in operating systems correspond to data types as they appear in some more recent programming languages. The object model seems, in some sense, to capture fundamental properties that pervade all aspects of modern operating systems: protection, naming, binding, data, procedures, and physical devices. A model of this nature seems to be necessary in order to realistically consider the validation of important properties of an operating system, such as correctness and reliability.

There are a substantial number of major problems that affect the entire fiber of the more advanced operating systems. Most of these problems appear in the newer system organizations, such as, data base operating systems, distributed systems, and networks of computers. In these new settings the problems tend to be an order of magni-